Claims

We claim:

1.   A method of obtaining the metadata of a data source, comprising:

creating a data repository having an entity structure which defines the metadata characteristics of a generic model data source;

5   accessing the data source to determine its construct;

configuring the data repository entities to reflect the construct of the data source; and

analyzing the data source in response to the configured data repository entities to obtain the source metadata.

10

2.   The method of claim 1, wherein said entity structure of said data repository is independent of the construct of the data source.

3.   The method of claim 1, wherein said entity structure of said data repository comprises plural entities, each of said plural entities corresponding to a different aspect of a generic model data source, and each having atomic elements corresponding to the metadata attributes of a generic entity, whereby

5   said plural entities collectively define the metadata characteristics of a generic model data source.

4.   The method of claim 3, wherein said step of analyzing includes:

obtaining, from the data source, those elements of source data which correspond to the metadata attributes of the configured data repository entities; and

5   recording the obtained elements of source data in said data repository, each in association with their corresponding metadata attribute.

5.   The method of claim 4, wherein said step of analyzing further includes:

inferring selected aspects of the data source structure on the basis of said recorded elements of source data, and recording the inferred data source aspects in said data repository;

5   presenting said inferred data source aspects for review by an operator; and

modifying said inferred data source aspects in response to command of the operator.

6.   The method of claim 4, wherein said data repository entities include at least one each of a database entity, a table entity, and a column entity, each thereof having plural metadata attributes, and each thereof having an associated child property entity, each said child property entity being capable of receiving

5   therein additional metadata attributes to be associated with its parent entity, as provided by an operator, thereby permitting modification of the parent entity attributes without altering said data repository entity structure.

7.   The method of claim 6, wherein:

said step of obtaining includes retrieving, as available from the data source, those elements of source data corresponding to the attributes of each of said database entity, said table entity, and said column entity.

5

8.   The method of claim 7, wherein said step of obtaining further includes:

recording, in a first area of said data repository, those elements of source data received in said step of retrieving, and inferring therefrom the data source schema .

5

9.   The method of claim 3 wherein:

said data repository entities include at least one each of a database analysis entity, a table analysis entity, and a column analysis entity; and wherein said step of analyzing includes,

acquiring, by elemental examination of the data source schema, those elements of source data corresponding to the metadata attributes of one or all of said database analysis entity, said table analysis entity, and said column analysis entity;

recording, in a second area of said data repository, those elements of source data received in said step of acquiring, and inferring therefrom the data source schema.

10. The method of claim 9, further comprising:

presenting said inferred data source schema for review by an operator; and

modifying said inferred data source schema in response to command of the operator.

11. The method of claim 10 wherein said elemental examination includes:

performing the step of acquiring the source data in an alternating sequence of receiving source data for an analysis entity, making inferences based on the received source data for that entity, and permitting operator modification of the inferences made for that entity prior to receiving source data for a following data analysis entity, whereby the inferences made in each interval are antecedent to the inferences made in succeeding intervals.

12. The method of claim 11 wherein said alternating sequence is ordered to first receive those elements of source data associated with said column analysis entity.

13. The method of claim 12 wherein the elements of source data associated with said table analysis entity are received following operator review of the inferences made in connection with said column analysis entity.

14.   Apparatus for obtaining the metadata of a data source, comprising:

a graphical user interface (GUI), responsive to commands entered by an operator; and

an analysis signal processor, responsive to said GUI, and adapted for connectivity to the data source, said analysis signal processor having a memory medium for storing signals, including program signals; said analysis signal processor, in response to said program signals:

creating within said memory medium a data repository having an entity structure which defines the metadata characteristics of a generic model data source;

accessing the data source to determine its construct;

configuring the data repository entities to reflect the construct of the data source; and

analyzing the data source in response to the configured data repository entities to obtain the source metadata.

15.   The apparatus of claim 14, wherein said entity structure of said data repository is independent of the construct of the data source.

16.   The apparatus of claim 14, wherein said entity structure of said data repository comprises plural entities, each of said plural entities corresponding to a different aspect of a generic model data source, and each having atomic elements corresponding to the metadata attributes of a generic entity, whereby said plural entities collectively define the metadata characteristics of a generic model data source.

17.   The apparatus of claim 16, wherein said step of analyzing includes:

obtaining, from the data source, those elements of source data which correspond to the metadata attributes of the configured data repository entities; and

46

5   recording the obtained elements of source data in said data repository,

each in association with their corresponding metadata attribute.

18.   The apparatus of claim 17, wherein said step of analyzing further

includes:

inferring selected aspects of the data source structure on the basis of said

recorded elements of source data, and recording the inferred data source aspects

5   in said data repository;

presenting said inferred data source aspects for review by an operator;

and

modifying said inferred data source aspects in response to command of

the operator.

10

19.   The apparatus of claim 17, wherein said data repository entities include

at least one each of a database entity, a table entity, and a column entity, each

thereof having plural metadata attributes, and each thereof having an associated

child property entity, each said child property entity being capable of receiving

5   therein additional metadata attributes to be associated with its parent entity, as

provided by an operator, thereby permitting modification of the parent entity

attributes without altering said data repository entity structure.

20.   The apparatus of claim 19, wherein:

said step of obtaining includes retrieving, as available from the data

source, those elements of source data corresponding to the attributes of each of

said database entity, said table entity, and said column entity.

5

21.   The apparatus of claim 20, wherein said step of obtaining further

includes:

recording, in a first area of said data repository, those elements of source

data received in said step of retrieving, and inferring therefrom the data source

5   schema .

22.   The apparatus of claim 16 wherein:

said data repository entities include at least one each of a database

analysis entity, a table analysis entity, and a column analysis entity; and wherein

said step of analyzing includes,

5   acquiring, by elemental examination of the data source schema, those

elements of source data corresponding to the metadata attributes of one or all of

said database analysis entity, said table analysis entity, and said column analysis

entity;

recording, in a second area of said data repository, those elements of

10   source data received in said step of acquiring, and inferring therefrom the data

source schema .

23.   The apparatus of claim 22, further comprising:

presenting said inferred data source schema for review by an operator;

and

modifying said inferred data source schema in response to command of

5   the operator.,

24.   The apparatus of claim 22 wherein said elemental examination includes:

performing the step of acquiring the source data in an alternating

sequence of receiving source data for an analysis entity, making inferences

based on the received source data for that entity, and permitting operator

5   modification of the inferences made for that entity prior to receiving source data

for a following data analysis entity, whereby the inferences made in each

interval are antecedent to the inferences made in succeeding intervals.

48

25.   The apparatus of claim 24 wherein said alternating sequence is ordered to first receive those elements of source data associated with said column analysis entity.

26.   The apparatus of claim 25 wherein the elements of source data associated with said table analysis entity are received following operator review of the inferences made in connection with said column analysis entity.

27.   A memory medium, for storing program signals to be used in controlling the operation of one or more signal processors and associated signal memory, in determining the metadata of a data source, the program signals controlling the signal processor in:

5   creating a data repository having an entity structure which defines the metadata characteristics of a generic model data source;

accessing the data source to determine its construct;

configuring the data repository entities to reflect the construct of the data source; and

10   analyzing the data source in response to the configured data repository entities to obtain the source metadata.

28.   The memory medium of claim 27, wherein said entity structure of said data repository is independent of the construct of the data source.

29.   The memory medium of claim 27, wherein said entity structure of said data repository comprises plural entities, each of said plural entities corresponding to a different aspect of a generic model data source, and each having atomic elements corresponding to the metadata attributes of a generic

5   entity, whereby said plural entities collectively define the metadata characteristics of a generic model data source.

30.   The memory medium of claim 29, wherein said step of analyzing includes:

obtaining, from the data source, those elements of source data which correspond to the metadata attributes of the configured data repository entities; and

recording the obtained elements of source data in said data repository, each in association with their corresponding metadata attribute.

31.   The memory medium of claim 30, wherein said step of analyzing further includes:

inferring selected aspects of the data source structure on the basis of said recorded elements of source data, and recording the inferred data source aspects in said data repository;

presenting said inferred data source aspects for review by an operator; and

modifying said inferred data source aspects in response to command of the operator.

32.   The memory medium of claim 30, wherein said data repository entities include at least one each of a database entity, a table entity, and a column entity, each thereof having plural metadata attributes, and each thereof having an associated child property entity, each said child property entity being capable of receiving therein additional metadata attributes to be associated with its parent entity, as provided by an operator, thereby permitting modification of the parent entity attributes without altering said data repository entity structure.

33.   The memory medium of claim 32, wherein:

said step of obtaining includes retrieving, as available from the data source, those elements of source data corresponding to the attributes of each of said database entity, said table entity, and said column entity.

34.   The memory medium of claim 33, wherein said step of obtaining further includes:

recording, in a first area of said data repository, those elements of source data received in said step of retrieving, and inferring therefrom the data source schema .

35.   The memory medium of claim 29 wherein:

said data repository entities include at least one each of a database analysis entity, a table analysis entity, and a column analysis entity; and wherein said step of analyzing includes,

acquiring, by elemental examination of the data source schema, those elements of source data corresponding to the metadata attributes of one or all of said database analysis entity, said table analysis entity, and said column analysis entity;

recording, in a second area of said data repository, those elements of source data received in said step of acquiring, and inferring therefrom the data source schema .

36.   The memory medium of claim 35, further comprising:

presenting said inferred data source schema for review by an operator;
and

modifying said inferred data source schema in response to command of
the operator.

37.   The memory medium of claim 36, wherein said elemental examination
includes:

performing the step of acquiring the source data in an alternating
sequence of receiving source data for an analysis entity, making inferences
based on the received source data for that entity, and permitting operator
modification of the inferences made for that entity prior to receiving source data
for a following data analysis entity, whereby the inferences made in each
interval are antecedent to the inferences made in succeeding intervals.

38.   The memory medium of claim 37, wherein said alternating sequence is
ordered to first receive those elements of source data associated with said
column analysis entity.

39.   The memory medium of claim 38, wherein the elements of source data
associated with said table analysis entity are received following operator review
of the inferences made in connection with said column analysis entity.